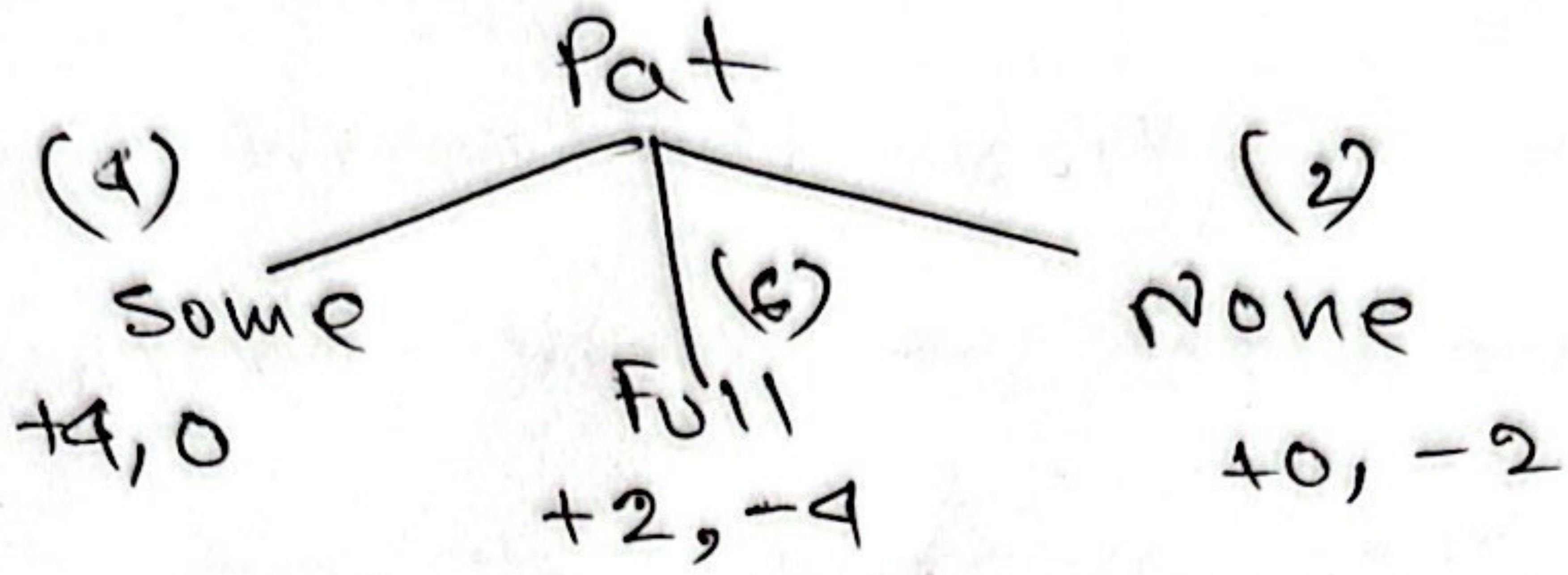
CT-1

1| Information gain for Pat attribute




Before using Pat

+6, -6

$$I_1 = -\frac{6}{12}\log_2\frac{6}{12} - \frac{6}{12}\log_2\frac{6}{12}$$

$$= 1$$

After using Pat

$$I_2 = \frac{4}{12}\times\left(\cancel{\log} - \frac{4}{4}\log_2\frac{4}{4} - \frac{0}{4}\log_2\frac{0}{4}\right)$$

$$+ \frac{6}{12}\times\left(-\frac{2}{6}\log_2\frac{2}{6} - \frac{4}{6}\log_2\frac{4}{6}\right)$$

$$+ \frac{2}{12}\times\left(-\frac{0}{2}\log_2\frac{0}{2} - \frac{2}{2}\log_2\frac{2}{2}\right)$$

$$= \frac{4}{12}\times 0 + \frac{6}{12}\times 0.92 + \frac{2}{12}\times 0$$

$$= 0.46$$

$\therefore$ Gain $= I_1 - I_2 = 1 - 0.46 = 0.54$

2) Neen of pruning a decision tree

- A tree migut have complex and many branches. but it may not generalize well to new or unseen data. So, to improve it's generalization ability, we nee to prune.

- To improve it's model complexity we need to prune a decision tree

- Pruning can also help to reduce the computational resources

- After ~~de~~ pruining, it may save our time to make a decision from the tree.

3) Given, $\sigma = 10$, $\mu = 40$, $x = 45$

$$P(x) = \frac{1}{\sqrt{2\pi} \times \sigma} \times e^{-\frac{(x-\mu)^2}{2 \times \sigma^2}}$$

$$= \frac{1}{\sqrt{2\pi} \times 10} \times e^{-\frac{(45-40)^2}{2 \times 10^2}}$$

$$= 0.0352$$ Ans

# Mid

~~Mid~~

(1) $$P(c|\underline{x}) = \frac{P(\underline{x}|c) \times P(c)}{P(\underline{x})}$$

O = Sunny, T = Cool, H = High, W = Weak

$$P(Yes|S,C,H,W) = \frac{P(S,C,H,W|Yes)}{P(S,C,H,W)}$$

$$P(No|S,C,H,W) = \frac{P(S,C,H,W|No)}{P(S,C,H,W)}$$

No = 5 total = 14

~~For P(Ye~~ Yes = 9

For Yes Class

$$P(S,C,H,W|Yes) = P(S|Yes) \times P(C|Yes) \times P(H|Yes) \times P(W|Yes) \times P(Yes)$$

$$= \frac{2}{9} \times \frac{3}{9} \times \frac{3}{9} \times \frac{5}{9} \times \frac{9}{14}$$

$= \cancel{0.01058}$

$= 0.007055$

$$P(S,C,H,W \mid No) = P(S \mid No) \times P(C \mid No) \times P(H \mid No) \times P(W \mid No) \times P(No)$$

$$= \frac{3}{8} \times \frac{1}{8} \times \frac{4}{5} \times \frac{1}{7} \times \frac{5}{14}$$

$$= 0.0049$$

<u>wind</u> (No)

(9) strong = $\cancel{5}$ = $\frac{5}{5}$

(5) weak = 0 = $\frac{0}{5}$

↓

strong = 5 + 1 = 6 = $\frac{6}{7}$

weak = 0 + 1 = 1 = $\frac{1}{7}$

AS 0.007055 > 0.0049

So, result is Yes.

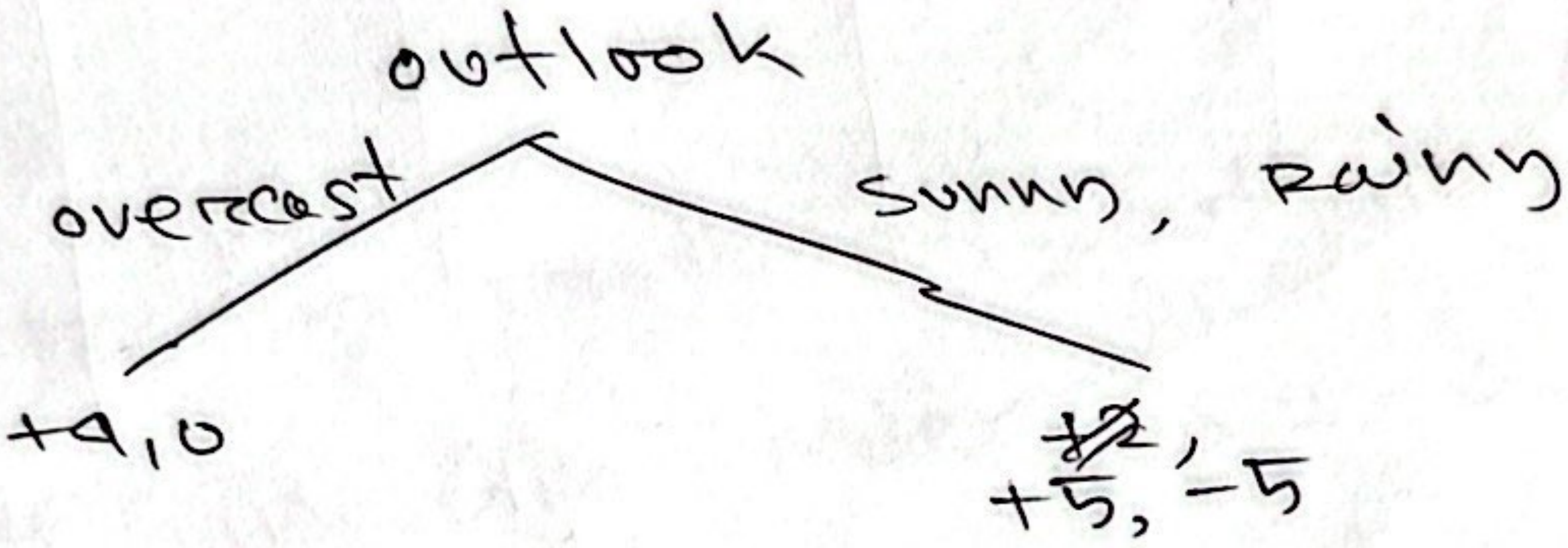
2) Gini index of outlook

Before using outlook

$$I_1 = 1 - \left(\frac{9}{14}\right)^2 - \left(\frac{5}{14}\right)^2$$

$$= 0.46$$

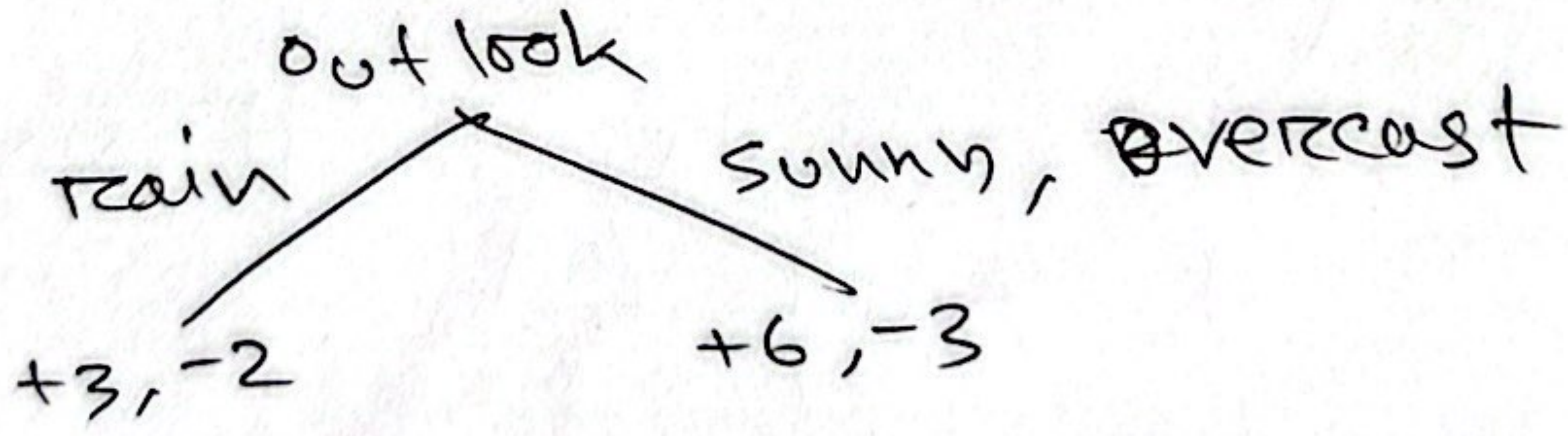
<u>After using outlook</u>




$$I_2 = \frac{4}{14}\left\{1 - \left(\frac{4}{4}\right)^2 - \left(\frac{0}{4}\right)^2\right\} + \frac{10}{14}\left\{1 - \left(\frac{5}{10}\right)^2 - \left(\frac{5}{10}\right)^2\right\}$$

$$= 0.36$$

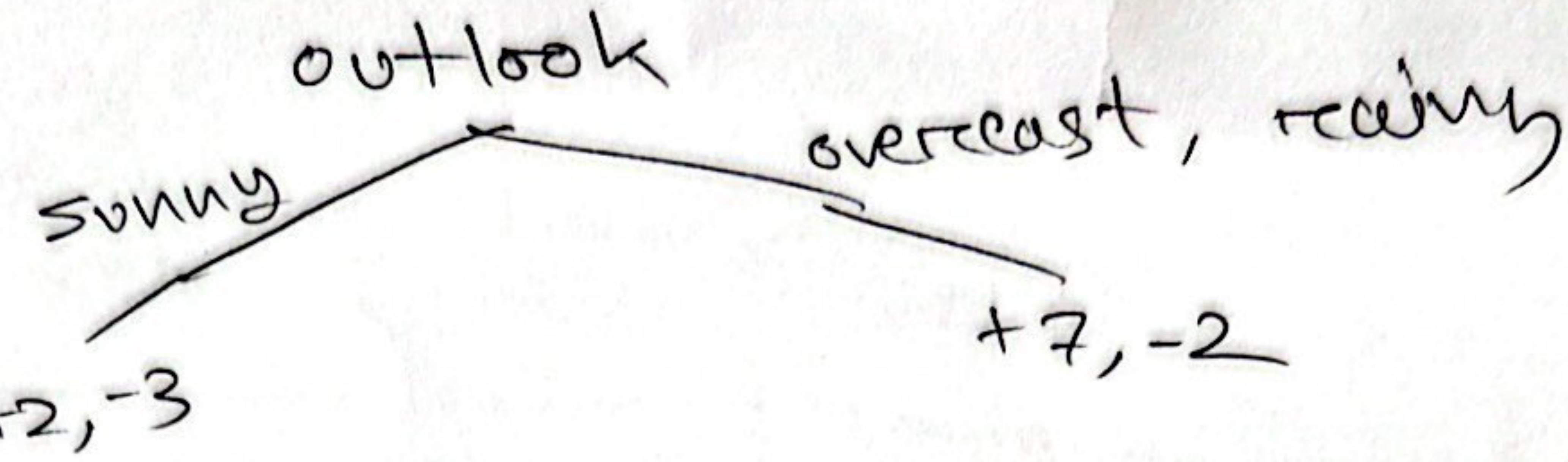
$\therefore$ change in Gini index $= 0.46 - 0.36$
$= 0.10$




$$I_2 = \frac{5}{14}\left\{1 - \left(\frac{3}{5}\right)^2 - \left(\frac{2}{5}\right)^2\right\} + \frac{9}{14}\left\{1 - \left(\frac{6}{9}\right)^2 - \left(\frac{3}{9}\right)^2\right\}$$

$$= 0.46$$

Change in Gini index $= 0.46 - 0.46 = 0.$

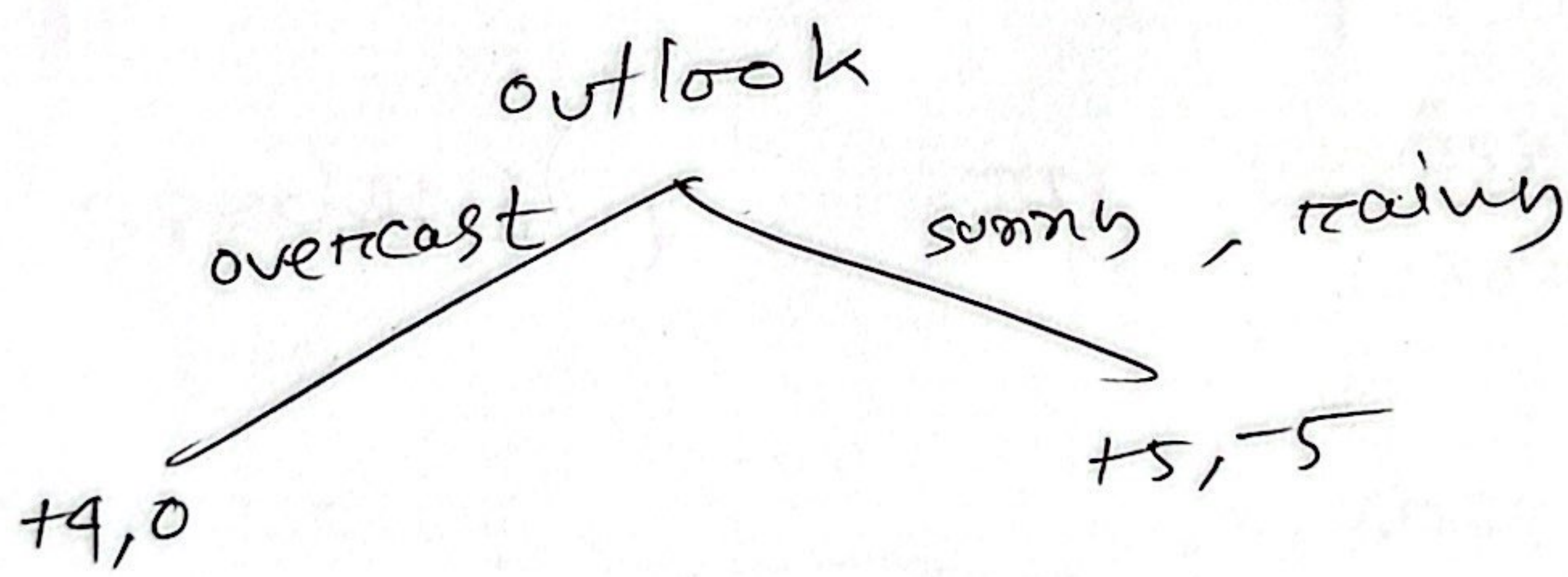



$$I_2 = \frac{5}{14}\left\{1-\left(\frac{3}{5}\right)^2-\left(\frac{2}{5}\right)^2\right\}+\frac{9}{14}\left\{1-\left(\frac{7}{9}\right)^2-\left(\frac{2}{9}\right)^2\right\}$$

$= 0.39$

Change in Gini Index $= 0.46 - 0.39$
$= 0.07$

so, ~~for~~ the root of outlook the change in Gini index is = 0.36



<u>Define</u>

TP FN

| | | Yes | No | |
|---|---|---|---|---|
| actual class | Yes | 6954 | 46 | 7000 |
| | NO | 412 | 2588 | 3000 |
| | | 7366 | 2634 | 10000 |

FP TN

1) accuracy $= \dfrac{6954 + 2588}{10000}$

2) error $=$ 1 − accuracy

3) sensitivity $= \dfrac{TP}{P} = \dfrac{6954}{7000}$

4) specificity $= \dfrac{TN}{N} = \dfrac{2588}{3000}$

5) precision $= \dfrac{TP}{TP + FP} = \dfrac{6954}{6954 + 412}$

6) recall $= \dfrac{TP}{TP+FN} = \dfrac{TP}{P} = \dfrac{6959}{7000}$

~~7) F = 2× Pr~~

3) Rule   Class   function

| Rule | Class | function |
|---|---|---|
| outlook = Sunny | NO | $D1^{+}, D2^{-}, D8^{-}$ $D9^{+}, D11^{+}$ |

accuracy $= \frac{3}{5} \times 100\%$

$= 60\%$

No = 3
Yes = 2
Total = 5

Coverage $= \frac{5}{14} \times 100\%$

$= 35.71\%$

$\left.D_1^{-} \; D_2^{-} \; D_8^{-}\right]$ TN

$\left.D_9^{+} \; D_{11}^{+}\right]$ FN

FP (False positive) = 0 Pu

FN (False Negative) = 2 Pu

Rule:

i) outlook = Sunny

~~+2~~,

Total
+ = 9
− = 5

after apply rules
(+2, −3) ~~কিন্তু~~ এখানে class = NO

so, +3, −2

Foil-Gain $= 3\left(\log_2 \frac{3}{3+2} - \log_2 \frac{5}{9+5}\right)$

$= \cancel{1.0.35}$

$= 2.25$

As Negative class
so, +5, −9

(+9, -5) +5, ~~+9~~

outlook = sunny , Humidity = High   Class = NO

$+\cancel{3}, -2$   $+3, 0$

$F.G_1 = -0.3$   $F.G2 = ?$

$$F.G2 = 3\left(\log_2 \frac{3}{3+0} - \log_2 \frac{3}{3+2}\right)$$

$= \cancel{3X}\ 2.211$ Ans

Scanned with CamScanner

4) $\overline{err}(M_1) = (30.5 + 32.2 + 20.7 + 20.6 + 31.0 + 41.0 + 27.7 + 26.0 + 21.5 + 26.0) / 10$

$= 27.72$

$\overline{err}(M_2) = (22.4 + 14.5 + 22.4 + 19.6 + 20.7 + 20.4 + 22.2 + 19.4 + 16.2 + 35.0 + \cancel{22.4}) / 10$

$= \cancel{19.03} \quad 21.28$

$(\overline{err}(M_1) - \overline{err}(M_2)) = 27.72 - \cancel{19.03}$ $21.28$

$= \cancel{8.69}$

$= 6.44$

$err(M_1) - err(M_2) = 30.5 - 22.4 = 8.1$

$32.2 - 14.5 = 17.7$

$20.7 - 22.4 = -1.7$

$20.6 - 19.6 = 1$

$31.0 - 20.7 = 10.3$

$41.0 - 20.4 = 20.3$

$27.7 - 22.2 = 5.5$

$26.0 - 19.4 = 6.6$

$21.5 - 16.2 = 5.3$

$26.0 - 35.0 = -9$

$\text{var}(M_1 - M_2)$

$$= \frac{1}{k} \sum \left[ \left( \text{err}(M_1)_i - \text{err}(M_2)_i \right) - \left( \overline{\text{err}}\, M_1 - \overline{\text{err}}\, M_2 \right) \right]^2$$

$$= \frac{1}{k} \Big[ (8.1 - 6.44)^2 + (17.7 - 6.44)^2 + (-1.7 - 6.44)^2$$
$$+ (1 - 6.44)^2 + (10.3 - 6.44)^2 + (20.3 - 6.44)^2$$
$$+ (5.5 - 6.44)^2 + (6.6 - 6.44)^2$$
$$+ (5.3 - 6.44)^2 + (-9 - 6.44)^2 \Big]$$

$$= \frac{1}{10} \times 673.31$$

$$= 67.331$$

$$t = \frac{\overline{\text{err}}(M_1) - \overline{\text{err}}(M_2)}{\sqrt{\dfrac{\text{var}(M_1 - M_2)}{k}}}$$

$$= \frac{6.44}{\sqrt{\dfrac{67.331}{10}}}$$

$$= 2.482$$

for 9 degrees of freedom (10 round - 1)

and 5% significance level,

the critical value, $z = 1.26$

$\therefore t = 2.481 > z = 1.26$

so, null hypothesis is rejected,

so, M1 and M2 has difference

and one is better.

As $\overline{err}(M_1) > \overline{err}(M_2)$

so M2 is better

5)

Total ~~exa~~example = 100

every example has equal weight =

$$\therefore \; w_i = \frac{1}{100} = 0.01$$

Given

accuracy = 90%

∴ ~~this~~ error = 10%

∴ error rate = 0.1

∴ update weight for correctly classified example is

$$\text{new, } W_i = \frac{\text{error}(w_i)}{1 - \text{error}(w_i)} \times w_i$$

$$= \frac{0.1}{1-0.1} \times \frac{1}{100}$$

$$= 0.001$$

Normalization

(1) new weight for correctly classified

$$= 0.001 \times \frac{1}{0.001 \times 90 + 0.01 \times 10}$$

$$= 0.001 \times \frac{1}{0.19} = 0.0053$$

(2) new weight for misclassified.

$$= 0.01 \times \frac{1}{0.001 \times 90 + 0.01 \times 10}$$

$$= 0.01 \times \frac{1}{0.19} = 0.053$$

(3) weight of the classifier

$$= \log\left(\frac{1 - error(M_i)}{error(M_i)}\right)$$

$$= \log\left(\frac{1 - 0.1}{0.1}\right)$$

$$= 0.954$$

6) OR function

Linear equation : $0.5x_1 + 0.5x_2 - 1.25 = 0$

Learning rate, $\alpha = 0.01$

hard threshold function

new weight = old weight + $\alpha$ (Target − ~~prediction~~ $h(x)$) × input

As hard threshold

$\therefore$ $0.5x_1 + 0.5x_2 - 1.25 \geq 0$ then output = 1

$0.5x_1 + 0.5x_2 - 1.25 < 0$ then output = 0

OR

| | $x_1$ | $x_2$ | Y |
|---|---|---|---|
| 1) | 0 | 0 | 0 |
| 2) | 0 | 1 | 1 |
| 3) | 1 | 0 | 1 |
| 4) | 1 | 1 | 1 |

for 1st input

input = 0, 0

output = 0 [target]

~~prediction~~ = $h(x) = 0$

$0.5 \times 0 + 0.5 \times 0 - 1.25$

$= -1.25 < 0$

so, $h(x) = 0$

No update

For 2nd input

input = 0, 1
output = 1 (target)

$0 \times 0.5 \times 0.5 - 1.25 = -0.75 < 0$

$\therefore h(x) = 0$ [ don't match with target ]
So update Needed

$w_1 = 0.5 + 0.01(1-0) \times 0 = 0.5$

$w_2 = 0.5 + 0.01(1-0) \times 1 = 0.51$

$w_0 = -1.25 + 0.01(1-0) \times 1 = -1.24$

equation: $0.5x_1 + 0.51x_2 - 1.24 = 0$

For 3rd input

input = 1, 0
output = 1 (target)

$0.5 \times 1 + 0.51 \times 0 - 1.24 = -0.74 < 0$

$\therefore h(x) = 0$ [ don't match with target
update needed

$w_1 = \cancel{0.5 \times 1 + 0.51}\ 0.5 + 0.01(1-0) \times 1 = 0.51$

$w_2 = 0.51 + 0.01(1-0) \times 0 = 0.51$

$w_0 = -1.24 + 0.01(1-0) \times 1 = -1.23$

so, equation: $0.51x_1 + 0.51x_2 - 1.23 = 0$

4th input

input = 1, 1

output = 1 [target]

$0.51 \times 1 + 0.51 \times 1 - 1.23 = -0.21 < 0$

$h(x) = 0$ [don't match with target]

update needed

$w_1 = 0.51 + 0.01(1-0) \times 1 = 0.52$

$w_2 = 0.51 + 0.01(1-0) \times 1 = 0.52$

$w_0 = -1.23 + 0.01(1-0) \times 1 = -1.22$

equation: $0.52x_1 + 0.52x_2 - 1.22 = 0$

31 derivation of sigmoid Function

$$\Delta w_i \propto \frac{\delta\, Loss}{\delta w_i}$$

$$Loss = (y - h_{\underline{w}}(x))^2$$

$$h_{\underline{w}}\underline{x} = \frac{1}{1+e^{-\underline{w}.\underline{x}}}$$

$$= \frac{\partial}{\partial w_i}(y - h_{\underline{w}}(x))^2$$

$$= 2(y - h_{\underline{w}}(x))\left[\frac{\partial}{\partial w_i}(y - h_{\underline{w}}(x))\right]$$

$$= 2(y - h_{\underline{w}}(x))\left[0 - \frac{\delta}{\delta w_i} h_{\underline{w}}(x)\right]$$

$$= -2(y - h_{\underline{w}}(x)) \times \frac{\delta}{\delta w_i}\left(\frac{1}{1+e^{-\underline{w}.x}}\right)$$

$$= -2(y - h_{\underline{w}}(x)) \times \frac{1}{1+e^{-\underline{w}.x}} \times \left(1 - \frac{1}{1+e^{-\underline{w}.x}}\right) \times \frac{\delta\, \underline{w}.x}{\delta w_1}$$

$$= -2(y - h\,\underline{w}(x)) \times h.\underline{w}(x) \times (1 - h.\underline{w}(x)) \times x_i$$

$$= \alpha(y - h.\underline{w}(x)) \times h.\underline{w}(x) \times (1 - h.\underline{w}(x)) \times x_i$$